Tanzania Journal of Science 48(2): 394-401, 2022
ISSN 0856-1761, e-ISSN 2507-7961

Modelling Internet Traffic Streams with Ga/M/1/K Queuing Systems under Self-similarity

Jumoke Popoola[1*], **Olusogo J Popoola**[2] **and Oyebayo R Olaniran**[1]
[1]*Department of Statistics, University of Ilorin, P.M.B. 1515, Ilorin, Nigeria*
[2]*Department of Computer Engineering, University of Ilorin, P.M.B. 1515, Ilorin, Nigeria*
**Corresponding author, e-mail: jmkbalpop@unilorin.edu.ng*
Other E-mails: olusogo@unilorin.edu.ng, olaniran.or@unilorin.edu.ng
Received 19 Jan 2022, Revised 23 May 2022, Accepted 30 May 2022, Published Jun 2022
DOI: https://dx.doi.org/10.4314/tjs.v48i2.14

Abstract

High-intensity concurrent arrivals of request packets in Internet traffic can cause dependence of event-to-event-times of the requests being served, which causes non-memoryless, modelled with heavy-tail distributions unlike common known traffics. The performance of Internet traffic can be examined using analytical models for the purpose of optimizing the system to reduce its operating costs. Therefore, our study examined a Ga/M/1/K Internet queue class (Gamma arrival processes, Ga; with memoryless-Poisson service process, M; a single server, 1, and K waiting room) and proposed specific derivations of its performance indicators. Real-life data of a corporate organisation Internet server was monitored at both peak and off-peak periods of its usage for Internet traffic data analysis. The minimum '0' in the arrival process indicates self-similarity and was assessed using Hurst parameter, H, and their (standard deviation). 'H' > 0.5 arrival process in the peak period only, indicates self-similarity. Performance of Ga/M/1/K was compared with various queuing Internet traffic models used in existing literatures. Results showed that the value of the waiting room size for Ga/M/1/K has closest ties with true self-similar model at peak-periods usage of the Internet, which indicates possible concurrent arrival of clients' requests leading to more usage of the waiting room, but with light-tailed queue model at the off-peak periods. Therefore, the proposed Ga/M/1/K model can assist in evaluating the performance of high-intensity self-similar Internet traffic.

Keywords: Internet traffic, self-similarity, Ga/M/1/K model, gamma distribution.

Introduction

Internet traffic streams under self-similarity

The Internet traffic queue has a peculiar nature of self-similarity, which makes it different from the traditional known queues experienced in common server systems. For instance, such other known traditional queues may occur when customers arrive in a banking hall and join a queue while waiting to be served; intending travellers arrive at a bus station and join the queue to purchase their bus tickets, etc. Rodrigues et al. (2022) worked on the assessment of health monitoring systems for elderly people, whereby chips were embedded inside aged people and connected to the Internet. The study assessed the M/M/c/K analytical queuing network model for the purpose of optimizing the performance of the system to minimize its costs of operations. Arrival processes of customers/clients' request to Internet servers in low intensified Internet traffics can be modelled by the Poisson distribution, M, and the service process modelled by the exponential distribution, M, (Alakiri et al. 2014, Olaniran and Abdullah 2020). In addition, the Memoryless property, M, in most traditional known queue systems, describes the independence of events, that is, event-to-event-times of service of customers. The M/M/1/K queue class model provides a model that can be used to explain such scenarios with its forgetfulness property and light-tailed distributions for describing such traffics.

On the contrary, the concurrent arrival of request packets to the server, particularly in a high-intensity Internet traffic is characterized by the self-similar property. This brings about the non-memoryless property of the Internet traffic steams with dependent event-to-event-times of service of customers and can be described by heavy-tailed distributions (Fras et al. 2013) but may be described by light-tailed distributions in a low Internet traffic stream (Teymori and Zhuang 2005).

However, queue class models that will be able to capture the self-similar property that is peculiar nature to the Internet traffic steams is extremely paramount. Furthermore, the specific derivations of the performance indicators of such Internet queue models will be able to provide appropriate estimates for evaluating the performance of the Internet server for its proper monitoring and management.

In some reviewed literatures, light-tailed distributions have been used to model Internet traffic; this is not adequate to describe the true self-similar property of the Internet traffic. In order to model the true self-similar property of the Internet, our current work proposes a Ga/M/1/K queue class model with a Gamma distribution, which is an appropriate heavy-tailed distribution for modelling concurrent arrivals of request packets with non-forgetfulness property. We derived specific performance indicators for our proposed Ga/M/1/K queue class model. This will provide appropriate estimates for understanding traffic behaviour in network aggregation points in light of self-similarity. For instance, this can give adequate estimate for predicting the queue lengths and waiting times in the models and can help maximize the use of limited resources in monitoring and managing the Internet.

Self-similarity in Internet traffic occurs when packets with the same burst length arrive at the same time or burst at the same inter-arrival interval on the server. Hurst was the first to identify self-similarity, which is now acknowledged to be present in many of the processes that explain natural and artificial occurrences. Popularity may be traced back to research that demonstrated the self-similar character of Internet flow and stressed the importance of understanding traffic behaviour in network aggregation points in light of self-similarity (Inácio et al. 2009). Self-similarity is commonly measured by estimating the Hurst parameter, H, (Rezakhah et al. 2012) where $0.5 < H < 1$ indicates whether or not a process is self-similar.

An object or process is defined as one that is self-similar if it is exactly or approximately similar to a portion of itself, that is, the whole has the same shape as one or more of the components. Self-similar processes can be described using heavy-tailed distributions (Fras et al. 2013). The main feature of heavy-tailed distributions is that they decay hyperbolically, rather than exponentially, as light-tailed distributions do. The Pareto distribution is the most basic heavy-tailed distribution currently available. According to Sheluhin et al. (2007), self-similar systems can also exhibit a feature of Long Range Dependence (LRD). Long Range Dependence describes the memory effect, in which a stochastic process's current value is heavily influenced by its previous values, and is explained by its autocorrelation function.

$0 < H < 0.5 \rightarrow$ SRD (Short Range Dependence and no similarity)

$0.5 < H < 1 \rightarrow$ LRD (Long Range Dependence and self-similarity)

The diffusion approximation was investigated by D'Arcy et al. (2014) as a solution to queuing analysis and approximations. Modelling of network and heavy traffic data systems, as well as issues resulting from the Internet and telecommunication traffic modelling using

Poisson models and assumptions were investigated in their study. The study's flaw is that it offered incorrect justifications for the Poisson arrival process's usefulness in capturing Internet data, in addition to network traffic. In Fras et al. (2013), a simulation tool for self-similar network traffic measurements, modelling, and simulations was developed. Only a few facts regarding self-similarity, long range interdependence, and probability were discussed in the work, which were used to define such stochastic processes. Moreover, (Jones 2004) focused on modelling self-similar networks by simulating self-similar processes. The study concentrated on the construction of an algorithm for generating and fitting self-similar processes, and merely demonstrated that the severity of self-similar processes is determined by the offspring distribution. According to Neame (2003), Gaussian models, especially LRD Gaussian models, are unable to accurately characterize modern Internet traffic; hence the study developed a new model called the Poisson Pareto burst process (PPBP). The model is a M/G/1 process with heavy-tailed on-off characteristics. The PPBP is shown to meet the basic requirements for a simple but accurate model of Internet traffic, with parameters that can fit observable data from an actual traffic stream and relationships between the PPBP and LRD Gaussian processes. These connections were investigated in the paper; however, the generated Internet data for the analysis is based on a Poisson process called PPBP, which assumes that current growth trends will become less accurate as models of core Internet traffic in the future.

According to Teymori and Zhuang (2005), two traffic flows were simulated, each with the same average rate and equal expected on and off durations. One is a non-heavy-tailed on-off source with exponentially distributed on and off durations. The study only found that the distribution for non-heavy-tailed traffic falls very quickly (exponentially), but not for heavy-tailed traffic, which has a relatively significant value even for a very long queue length, resulting in the potential of a very long delay and a high probability of packet loss. The study's flaw is that the Pareto distribution is fitted to the heavy-tailed source, but this does not depict the genuine scenario on the Internet at a given time when the traffic pattern may not be Pareto.

Materials and Methods

Some existing distributions for modelling Internet traffic

Exponential distribution: This happens to be the default model for monitoring Internet traffic before the advent of self-similar and Long Range Dependency issues. Exponential distribution belongs to the class of light-tailed distribution because of its memoryless property. The probability density function of exponential distribution given by Alakiri et al. (2014), and Olaniran and Abdullah (2020) is:

$$f(x,\theta) = \theta e^{-\theta x}, x > 0 \qquad (1)$$

and the cumulative distribution function (CDF) is:

$$F(x) = 1 - e^{-\theta x}$$

where; the θ is the intensity parameter of the process.

Erlang distribution: The Erlang distribution is a continuous probability distribution with wide applicability primarily due to its relation to the Exponential and Gamma distributions. The probability density function of the Erlang distribution is the same as the Gamma distribution except for the possible values to which the shape parameter k can assume. Thus, the PDF of Erlang distribution described by (Brockmayer, 1948) is;

$$f(x,k,\theta) = \frac{x^{k-1}\theta^k e^{-\theta x}}{(k-1)!}, k = 1,2,3, \ldots \; ; x > 0 \qquad (2)$$

and the cumulative distribution function (CDF) is:

$$F(x,k,\theta) = 1 - \sum_{i=0}^{k-1} \frac{e^{-\theta x}(\theta x)^i}{i!}, x > 0$$

Pareto distribution: The Pareto distribution has a long tail and follows a power law across its range. According to Alakiri et al. (2014), the PDF of the Pareto distribution is defined as:

$$f(x) = ka^k x^{-k-1}, 0 < a \leq x \qquad (3)$$

and the cumulative distribution function (CDF) is:

$$F(x) = 1 - (a/x)^k$$

where the constant a represents the smallest possible value of the random variable x, and k is the shape parameter of the distribution.

Gamma distribution: According to Tran-Gia et al. (2001), the Gamma distribution is also utilized in publications for simulating Internet traffic. The Gamma distribution's PDF is defined as follows:

$$f(x,k,\theta) = \frac{x^{k-1}\theta^k e^{-\theta x}}{\Gamma_k}, k > 0\,; x > 0 \quad (4)$$

and the cumulative distribution function (CDF) on the support of X is:

$$F(x,k,\theta) = \frac{\Gamma(\ k,\ \ \theta x\ )}{\Gamma_k}, k > 0\,; x >$$

where $\Gamma(s,x) = \int_0^x t^{s-1}e^{-t}dt$ For $s > 0\,; x > 0$ is the incomplete Gamma function, k is the shape and θ is the intensity parameter of the distribution; and Γ is the Gamma function which has the formula;

$$\Gamma(s) = \int_0^\infty t^{s-1}e^{-t}dt$$

for $s > 0$ is the Gamma function.

The proposed $G/M/1/K$ self-similar model for burst Internet traffic

The $G/M/1$ queue is a single-server queue, where the arrival process follows a general distribution and the service process has an Exponential distribution with mean service time $1/\mu$, that is

$$B(x) = 1 - e^{-\mu x},\ x > 0] \quad (5)$$

while the arrival process is general with mean inter-arrival time equal to the mean of inter-arrival time of the distribution G, in G/M/1/K . Request-packets arrive one at a time, and their inter-arrival delays are dispersed equally and independently.

Review and proposed specific derivatives of parameter formulae for evaluating performance indicators for arrival and transmission of packets in a $G/M/1$ traffic

For the claim of the parameter formula for the performance indicators, Stewart (2009) gave the stationary distribution p_i of packets in a $G/M/1$ traffic as;

$$p_i = \rho(1-\xi)\xi^{i-1},\ i > 0 \quad (6)$$

where ξ is interpreted as the traffic intensity and it is computed from the relation given below.

$$\xi = F_A^*(\mu - \mu\xi) \quad (7)$$

where $F_A^*(\mu - \mu\xi)$ is the laplace transform of the arrival process evaluated at $s = \mu - \mu\xi$ and $0 < \xi < 1$.

G/M/1 where G is Gamma

Stewart (2009) used the Pollaczek Khintchine transform to state the stationary distribution of request-packets at the server end in a $G/M/1$ as;

$$p_i = \rho(1-\xi)\xi^{i-1},\ i > 0 \quad (8)$$

or

$$p_i = (1-\xi)\ \xi^i\ ,\ i > 0 \quad (9)$$

According to Pollaczek Khintchine transform, where, ξ is computed from $\xi = F_A^*(\mu - \mu\xi)$; and $F_A^*(\mu - \mu\xi)$ is the Laplace transform of the arrival process evaluated at $s = \mu - \mu\xi$ and $0 < \xi < 1$ since G in G/M/1/K follows Gamma distribution for the inter-arrival process.

Now, the proof of the parameter formula is given thus;

Suppose the inter-arrival time, t, distribution follows Gamma with parameters α and (λ), then the density function of t, is;

$$f(t,\alpha,\lambda) = \frac{t^{\alpha-1}\lambda^\alpha e^{-\lambda t}}{\Gamma(\alpha)} \quad \alpha; \lambda > 0;\ t > 0 \quad (10)$$

In this study, the specific derived Laplace transform of the arrival process, evaluated at $s = \mu - \mu\xi$ and $0 < \xi < 1$ for G/M/1/K queue model, when G in G/M/1/K follows a Gamma distribution for the inter-arrival process is;

$$L(t) = \left[\frac{\lambda}{s+\lambda}\right]^\alpha \quad (11)$$

According to Stewart (2009), since ξ is the Laplace transform of the arrival process, evaluated at $s = \mu - \mu\xi$ and $0 < \xi < 1$ for G/M/1/K queue model, which is computed from $\xi = F_A^*(\mu - \mu\xi)$;

then; specific derivation in this study for ξ in G/M/1/K when G is Gamma for the arrival process is;

$$\xi = F_A^*(\mu - \mu\xi) = \left[\frac{\lambda}{s+\lambda}\right]^\alpha; \quad (12)$$

That is,

$$\xi = F_A^*(\mu - \mu\xi) = \left[\frac{\lambda}{\mu(1-\xi)+\lambda}\right]^\alpha \quad (13)$$

$$\xi = \left[\frac{\lambda}{\mu(1-\xi)+\lambda}\right]^{\alpha} \quad (14)$$

$$\xi^{\frac{1}{\alpha}} = \left[\frac{\lambda}{\mu(1-\xi)+\lambda}\right] \quad (15)$$

After solving some algebra, the solutions to ξ are $\xi = 1$ and $= \frac{\lambda}{\alpha\mu}$.

Since it is required that $\xi < 1$, for a stationary process, then the required solution is $\frac{\lambda}{\alpha\mu}$

In this study, the proof of the specific derived traffic intensity in a $G/M/1$, when G is Gamma is, τ, where $\tau = \frac{\lambda}{\alpha\mu}$.

According to Pollaczek Khintchine transform for G/M/1/K, the probability of having no connection of multiple request-packets at the server end, denoted by, p_0, is $P(z)$ when $z = 0$ and similarly, the probability of having i connection of multiple request-packets at the server end is $p_i = \frac{d^i P(z)}{dz^i}$ evaluated at $z = 0$.

Using Pollaczek Khintchine transform equation, $p_i = (1-\xi)\,\xi^i$, $i > 0$, in this study, the proof of the specific derived distributions of request-packets at the server end for G/M/I/K when G is Gamma is;

$$p_i = \begin{cases} 1-\tau, & i = 0 \\ (1-\tau)\tau^i, & i > 0 \end{cases} \quad (16)$$

Using Pollaczek Khintchine transform equation, in this study, the specific derived probability of reaching the maximum blocking bandwidth available to the server for concurrent connections of request-packets, often interpreted as the blocking probability of the bandwidth available to the server, p_K, for G/M/I when G is Gamma is;

$$p_K = (1-\tau)\tau^K \quad (17)$$

In the same line, the waiting room size, K, given p_K is;

$$K = \frac{log\left[\frac{p_K}{(1-\tau)}\right]}{log(\tau)} \quad (18)$$

Using Pollaczek Khintchine formula, $L = L_q + \rho$; in this study, the proof of the specific derived distributions of request-packets on the queue at the server end for G/M/1/K when G is Gamma is;

$$p_i(L_q) = \begin{cases} 1-2\tau, & i = 0 \\ (1-\tau)\tau^i - \tau, & i > 0 \end{cases} \quad (19)$$

The derivation in equation (19) above indicates a close link with the $M/M/1$ queue; the mean number of request-packets at the server end for $M/M/1$ is;

$$L = \frac{\rho}{1-\rho}$$

where, $\rho = \frac{\lambda}{\mu}$.

Thus, by analogy, in this study, the specific derived mean number of request-packets at the server end for $G/M/1$ where the inter-arrival time distribution is Gamma is;

$$L = \frac{\tau}{1-\tau} \quad (20)$$

Similarly, in this study, the proof of the specific derived mean number of request-packets on the queue at the server end for $G/M/1$ where the inter-arrival time distribution is Gamma is; $L_q = L - \tau$

$$L_q = \frac{\tau}{1-\tau} - \tau \quad (21)$$

Since there exists a close association between $M/M/1$ and $G/M/1$ queue, the only difference is the utilization factor, which in $G/M/1$ queue is ξ. The distribution of time spent by request-packets at the server end in $M/M/1$ by literature is given as;

$$W(t) = \mu(1-\rho)e^{-[\mu(1-\rho)]t}, t > 0 \quad (22)$$

while the distribution of time spent by request-packets on the queue at the server end in M/M/1/K is;

$$W_q(t) = \rho\mu(1-\rho)e^{-[\mu(1-\rho)]t}, t > 0 \quad (23)$$

Therefore, by analogy, in this study, the specific derived distribution of time spent by request-packets at the server end for $G/M/1$ queue when G is Gamma is;

$$W(t) = \mu(1-\xi)e^{-[\mu(1-\xi)]t}, t > 0 \quad (24)$$

while in this study, the proof of the specific derived distribution of time spent by request-packets on the queue at the server end for G/M/1/K when G is Gamma is;

$$W_q(t) = \xi\mu(1-\xi)e^{-[\mu(1-\xi)]t}, t > 0 \quad (25)$$

Now, assuming Gamma distribution as the inter-arrival distribution, where it has been derived in this study that, $\xi = \tau = \frac{\lambda}{\alpha\mu}$, the corresponding distributions of time spent by request-packets at the server end and request-packets on the queue at the server end for G/M/1/K when G is Gamma, respectively, are;

$$W(t) = \mu[1-\tau]e^{-[\mu(1-\tau)]t}, t > 0 \quad (26)$$

and,

$$W_q(t) = \tau\mu[1-\tau]e^{-[\mu(1-\tau)]t}, t > 0 \quad (27)$$

It is obvious that equation (27) follows Exponential distribution with parameter, $\mu[1-\tau]$; mean times spent by request-packets at the server end and mean time spent request-packets on the queue at the server end for G/M/1/K when G is Gamma, respectively, are;

$$W = \frac{1}{\mu[1-\tau]} \quad (28)$$

and,

$$W_q = \frac{\tau}{\mu[1-\tau]} \quad (29)$$

Results and Discussion

Application to real life network traffic

The Transmission Control Protocol (TCP) for Internet computing of network flow of University of Ilorin, Nigeria with Internet Protocol address (www.unilorin.edu.ng) was monitored as sample network flow for this research. Packet arrival and transmission times were recorded in form of TCP connection format. A total of 1, 255, 981 clients' request-packets were observed between the period of April 19th, 2016 to April 20th, 2016. The data were further subdivided into peak and off-peak and the considered Internet traffic models in our study were fitted to the data. The off-peak period data cover the time period of 6:27 pm of April 19th, 2016 to 7:57 am of April 20th 2016. A total of 169, 628 request-packets were collected for the off-peak period analysis.

Similarly, the peak period data cover the time period of 8:27 am to 3:57 pm of April 20th, 2016. A total of 1,086,363 request-packets were collected for the peak period analysis. Table 1 shows the descriptive statistics of the real-life data for both arrival and transmission processes.

Table 1: Results of arrival and transmission processes of a Ga/M/1/K Internet real-life data

Period	Process	Number of observations	mean	Standard deviation	Minimum	Maximum
Peak	Arrival	1086363	2.6455	21.5006	0	1801
(8:27 am-3:57 pm)	Service transmission	1086363	1.4240	3.01307	0.000411	100.4988
Off peak	Arrival	169628	1.2196	249.0301	0.000240	86400
(6:27 pm-7.57 am)	Service transmission	169628	1.2194	2.4455	0.000391	100.6715

Time unit: Seconds.

Table 1 shows the descriptive summary in terms of mean, standard deviation, minimum and maximum. The minimum value 0 in the arrival process at the peak period indicates the possibility of concurrent arrivals in a Ga/M/1/K queue model, that is, some of the packets arrive on the network at the same time, with mean arrival process greater than that of the transmission, which suggests self-similarity in the Internet arrival process. The absence of this structure in the off-peak period transmission process indicates the possibility of a memoryless process. The validity of the data can be observed from the fact that irrespective of the partitioning, the maximum processing time is approximately 101 s. Also, the stability of the network can be observed from the mean of the inter-arrival time being greater than mean transmission time. Hence the network process monitoring is possible. The standard deviation estimates relatively larger than their means suggest adequacy of skewed distribution for the arrival and transmission process.

To further examine the self-similarity nature, Hurst index (H) was estimated using R/S statistics (Mandelbrot and Van Ness, 1968). The Hurst estimates results in Table 2 reveal that the arrival process at the peak period is more self-similar than the transmission process while at the off-peak period the self-similarity is low for both the arrival and transmission process. The peak and off-peak periods partitioning used here represent low and high traffic.

Table 2: Hurst index, H, estimate and their (standard error) for Ga/M/1/K real-life data

Period	Number of observations	Arrival process	Transmission process
Peak (8:27 am-3:57 pm)	1086363	H = 0.8636*** Standard error (0.2271)	H = 0.4786*** Standard error (0.1584)
Off peak (6:27 pm-7.57 am)	169628	H = 0.4022*** Standard error (0.1154)	H = 0.3814*** Standard error (0.0761)

*** Significant at 5% level.

Tables 3 and 4 present the modelling results for the peak and off-peak periods.

Table 3: Waiting room size, K, for various G/M/1/K models for the peak period data

Model	Waiting room size, K
Self-similar traffic (S/S/1/K)	9317
M/M/1/K	1204
Ga/M/1/K	9316
Ln/M/1/K	2763

Table 4: Results for waiting room size for the empirical true self-similar traffic model S/S/1/K and various G/M/1/K models for the off-peak period data

Model	Waiting room size, K
Self-similar traffic (S/S/1/K)	336
M/M/1/K	245
Ga/M/1/K	248
Ln/M/1/K	353

Conclusion

We created approximate performance metrics for a Ga/M/1/K class of Internet traffic model in this study, where ‘Ga’ is Gamma distribution. Various G/M/1/K Internet traffic model performances (‘G’ in the Kendall notation representing a general distribution of arrival process) were specifically assessed for Gamma, Ga, and Log-normal, Ln, arrival process distributions, as well as their ties to a M/M/1 Internet traffic model and Internet traffic model with true self-similar, ‘S’, arrival and transmission processes, S/S/1/K. Based on the findings from a real-life Internet traffic data of a corporate organisation, the minimum ‘0’ in the arrival process at the peak period of the Internet usage indicates possibility of concurrent arrivals leading to possibility of self-similarity, which is assessed using Hurst parameter, H, and their (standard deviation). ‘H’ > 0.5 arrival process in the peak period only, indicates self-similarity. Performance of Ga/M/1/K was compared with various queuing Internet traffic models used in existing literatures. Results show the waiting room size for Ga/M/1/K has closest ties with true self-similar model, S/S/1/K, which indicates largest size of the waiting room being utilized for concurrent arrival of clients' request in a possible bursty Internet traffic at peak-periods usage of the Internet, but with light-tailed queue model, M/M/1/K at the off-peak periods. Therefore, the proposed Ga/M/1/K model was found to be adequate and can assist in evaluating the performance of high-intensity self-similar Internet traffic.

To provide a final stand on modelling self-similar Internet traffic, the current work can be extended by comparing it to other classes of traffic models and heavy tailed distributions other than the distributions considered in our study.

Acknowledgement
This research was partly supported by the University of Ilorin, Ilorin, Nigeria.

Conflict of Interest: The authors declare no conflict of interest.

References

Alakiri OH, Oladeji A, Benjamin CB, Okolie CC and Okikiola MF 2014 The desirability of pareto distribution for modeling modern internet traffic characteristics. *Int. J. Novel Res. Eng. Appl. Sci.* 1(1): 1-9.

Brockmayer E 1948 Table of Erlang's loss formula, The life and works of AK Erlang. *Trans. Danish Acad. Tech. Sci*. 1: 268-275.

D'Arcy S, MacHale E, Seheult J, Holmes M, Hughes C, Sulaiman I and McCourt J 2014 A method to assess adherence in inhaler use through analysis of acoustic recordings of inhaler events. *PloS One* 6: e98701.

Fras M, Mohorko J and Čučej Ž 2013 Limitations of a Mapping Algorithm with Fragmentation Mimics (MAFM) when modeling statistical data sources based on measured packet network traffic. C*omputer Networks* 57(17): 3686-3700.

Inácio PR, Lakic B, Freire MM, Pereira M, and Monteiro PP 2009 The design and evaluation of the simple self-similar sequences generator. Information Sciences. *Inf. Sci.* 179(23): 4029-4045.

Jones OD 2004 Fast, efficient on-line simulation of self-similar processes. *Thinking in Patterns* 165-176.

Mandelbrot BB and Van Ness JW 1968 Fractional Brownian motions, fractional noises and applications. *SIAM Review* 10(4): 422-437.

Neame T 2003 *Characterisation and modelling of Internet traffic streams*. PhD thesis. Department of Electrical and Electronic Engineering, University of Melbourne.

Olaniran OR and Abdullah MA 2020 Subset selection in high-dimensional genomic data using hybrid variational Bayes and Bootstrap priors. In *Journal of Physics*: *Conference Series* 1489(1): 012030, *IOP Publishing*.

Rezakhah S, Philippe A and Modarresi N 2012 Estimation of scale and hurst parameters of semi-selfsimilar processes. *arXiv preprint arXiv* 1207.2450.

Rodrigues L, Rodrigues JJ, Serra AD and Silva FA 2022 A Queueing-based model performance evaluation for internet of people supported by fog computing. *Future Internet* 14(1): 23.

Sheluhin O, Smolskiy S and Osin A 2007 Self-similar processes in telecommunications. John Wiley & Sons.

Stewart WJ 2009 Probability, Markov chains, queues, and simulation: the mathematical basis of performance modeling. Princeton, New Jersey: Princeton University Press.

Teymori S and Zhuang W 2005 Queue analysis for wireless packet data traffic. In International Conference on Research in Networking (pp. 217-227). Berlin, Heidelberg. Springer.

Tran-Gia P, Staehle D and Leibnitz K 2001 Source traffic modeling of wireless applications. *AEU-International Journal of Electronics and Communications* 55(1): 27-36.